B1TSCOP1C

VA Selects Praedico Next Generation Big Data Biosurveillance Application

The Veterans Administration (VA) is one of the largest and most highly regarded health care organizations in the world. But its charter extends far beyond serving the medical needs of veterans. One of its primary missions is research. The VA was the first healthcare system to successfully implement an integrated electronic health record (EHR) system, which is used throughout the Veterans Health Administration (VHA) in all health care settings including inpatient, outpatient, and long-term care. With its high number of enrollees providing a rich source of medical data, the VHA has become a research

powerhouse and produces scientific advances that benefit all Americans, not just veterans.

One of VHA's research initiatives is the Healthcare Associated Infection and Influenza Surveillance System (HAIISS). HAIISS plays an important role in national security because it creates an infection identification, reporting and alert system, essentially a targeted biosurveillance system for a specific set of diseases and outbreaks. To expand its capabilities in this area, the VHA selected Praedico® Biosurveillance, a state of the art biosurveillance toolset designed

for early detection, monitoring, and forecasting of infectious disease outbreaks. Using Praedico Biosurveillance, the VHA will be at the forefront of biosurveillance capabilities and provide improved healthcare outcomes for veterans.

With access to a large amount of medical data from patients geographically distributed across the country, the VHA is uniquely positioned to monitor and predict outbreaks. For the past ten years, the VHA has been using a software application originally developed by the Department of Defense(DoD) that analyzes patient medical data to detect disease outbreaks and other public health threats. The system provides essential biosurveillance capabilities, but the VHA wanted to expand its capabilities beyond what the software could do. For example, the VHA has the mandate to integrate its data set with the DoD's and ultimately with various other federal, state, and local agencies and needed new tools that would allow them to share health data in real-time. Additionally, with the volume of medical data rapidly growing,

sometimes doubling as new data sets are integrated, the VA recognized that only way to avoid long processing times was to implement a new scalable computing platform. Finally, the VA was concerned about the high frequency of false positive alerts reported by the system. Every alert must be evaluated by analysts to determine whether it signals a true event or is instead a benign anomaly and this takes a great deal of time and effort. In fact, in some cases, the problem of false positives has led analysts to ignore alerts altogether and use the system to examine the data heuristically. Frequent false positives are an inherent consequence of systems that rely on statistical algorithms that detect events by looking for outliers or other anomalies. However, by adding machine learning algorithms to the mix, which would provide context awareness to event detection, the VHA could accurately filter out many false positives.

The heart of the VHA's biosurveillance programs is its huge repository of patient medical data. In general, as a data set becomes larger and more diverse, it strengthens the predictive capabilities of the system. Therefore, the integration of the VHA's data

with the DoD's was an opportunity to improve the quality of results for both organizations. The VHA deployed the Praedico Data Platform to extract and ingest data from VHA's advanced electronic health records (EHR) system and other sources, including the DoD's. The Praedico Data Platform processes the data, which includes cleansing, normalization, correlation, conflation and standardization and finally publishes the data to consuming applications and services within and outside the VHA. This established a comprehensive electronic surveillance data repository to monitor HAIISS events such as bug/drug resistance, healthcare associated infections, influenza, and other emerging infectious diseases or syndromes associated with natural and/or bioterrorist activity. Data is gathered and integrated multiple times throughout the day, resulting in near real-time analysis and reporting.

Before making the decision to adopt Praedico Biosurveillance, the VHA compared its performance and capabilities with its existing solution. Incorporating over 5 years of syndromic surveillance, 513 million VA patient records were

analyzed for identical syndromic surveillance groupings from outpatient ICD-9 diagnosis codes. This included training the Praedico machine learning layer using hundreds of false positive and true positive syndromic alerts. To guarantee high detection recall, the Praedico algorithm leverages many known detection algorithms, including versions of CDC, CUSUM, EWMA, and regression models.

The VHA's test showed that Praedico Biosurveillance provides the same level of analytical capabilities as its existing solution and requires significantly less data storage. A key capability provided by Praedico Biosurveillance was the ability to perform analysis on integrated data from both the VA and the DoD. Queries have reduced response times, often producing results in seconds for analyzing databases with millions of records.

With Praedico Biosurveillance now fully deployed, the VHA could also use the application to predict, detect and monitor other bio-related health events such as antibiotic resistance trends, bioterrorist events, intensive care unit devices, influenza outbreaks, surgical site infections, infectious diseases of public health significance.

BITSCOPIC